```
sau96I
                                                                   haeIII
                                                                   asuI
                                                                  sau96I
                                                                  nlaIV
                                                                  hgiJII
ox to springs
                                                                  eco01091
                                                                  bsp1286
                                                                  banII
                                                                  asul
                                                             aval apal mnil
                                      mn l I
                                                           mn l I
                                mbolI
                       GGA CTT GTC TTC CTC GTC CTG CTG TTC CTC GGG GCC CTC GGA CTG
                1
                       Gly Leu Val Phe Leu Val Leu Leu Phe Leu Gly Ala Leu Gly Leu
              -18
                               haeIII
                                                                      hinPI
                             eael
                                                                      hhaI
                             cfrI
               46 TGT CTG GCT GGC CGT AGG AGA AGG AGT GTT CAG TGG TGC GCC GTA TCC
               -3 Cys Leu Ala Gly Arg Arg Arg Ser Val Gln Trp Cys Ala Val Ser
                             haelll
                           nn l I
                       aval hael
               94 CAA CCC GAG GCC ACA AAA TGC TTC CAA TGG CAA AGG AAT ATG AGA AAA
               14 Gln Pro Glu Ala Thr Lys Cys Phe Gln Trp Gln Arg Asn Met Arg Lys
                                             fnu4HI
                               mn l I
                                                                 pleI
                                             bbv I
                           sau96I
                                                                 hinfl
                                                                                bsrI
                                            aluI
                           haelll
                                                                           fokI
                                                              bsma I
                                          pvuII
                           asul
              142 GTG CTG GGC CCT CCT GTC AGC TGC ATA AAG AGA GAC TCC CCC ATC CAG
               30 Val Arg Gly Pro Pro Val Ser Cys Ile Lys Arg Asp Ser Pro Ile Gin
                             haeIII
                            hael
                                                       haelll
                         scrFI
                                                      sau961
                         ecoRII
                                                      asuI
                                                             sfaN1
                         bstNI
              190 TGT ATC CAG GCC ATT GCG GAA AAC AGG GCC GAT GCT GTG ACC CTT GAT
               46 Cys Ile Gin Ala Ile Ala Glu Asn Arg Ala Asp Ala Val Thr Leu Asp
                                                       sau96I
                                                       nlaIV
                                                   scrF1
                                                   ecoRII
                                                   bstNI
                                                haelll
                                                       haelll
                                              stuI
                                                       asul
                                        mn l I
                                              haeI
              238 GGT GGT TTC ATA TAC GAG GCA GGC CTG GCC CCC TAC AAA CTG CGA CCT
               62 Gly Gly Phe Ile Tyr Glu Ala Gly Leu Ala Pro Tyr Lys Leu Arg Pro
```

FIG.-1A

sau96I avall asul accI nlaIV fnu4HI 286 GTA GCG GCG GAA GTC TAC GGG ACC GAA AGA CAG CCA CGA ACT CAC TAT 78 Val Ala Ala Glu Val Tyr Gly Thr Glu Arg Gln Pro Arg Thr His Tyr fnu4HI mboII bbv I aluI fnu4HI aluI hphI pvuII 334 TAT GCC GTG GCT GTG GTG AAG AAG GGC GGC AGC TTT CAG CTG AAC GAA 94 Tyr Ala Val Ala Val Val Lys Lys Gly Gly Ser Phe Gln Leu Asn Glu haelll sau96I stul avall bglI hael asul fokI 382 CTG CAA GGT CTG AAG TCC TGC CAC ACA GGC CTT CGC AGG ACC GCT GGA 110 Leu Gln Gly Leu Lys Ser Cys His Thr Gly Leu Arg Arg Thr Ala Gly sau961 avall asul nlaIV 430 TGG AAT GTC CCT ACA GGG ACA CTT CGT CCA TTC TTG AAT TGG ACG GGT 126 Trp Asn Val Pro Thr Gly Thr Leu Arg Pro Phe Leu Asn Trp Thr Gly aluIhgiJII fnu4HI bsp1286 dde I aluI banII bbvI dde I mn II pvuII mboll pvull 478 CCA CCT GAG CCC ATT GAG GCA GCT GTG CAG TTC TTC TCA GCC AGC TGT 142 Pro Pro Glu Pro Ile Glu Ala Ala Val Gln Phe Phe Ser Ala Ser Cys msp I hpall scrFI ncil caull 526 GTT CCC GGT GCA GAT AAA GGA CAG TTC CCC AAC CTG TGT CGC CTG TGT 158 Val Pro Gin Ala Asp Lys Gly Gin Phe Pro Asn Leu Cys Arg Leu Cys

> nlaIV scrFI ecoRII

mnlI bstNI

574 GCG GGG ACA GGG GAA AAC AAA TGT GCC TTC TCC CAG GAA CCG TAC 174 Ala Gly Thr Gly Glu Asn Lys Cys Ala Phe Ser Ser Gln Glu Pro Tyr

nlaIV hgiCI ddel bsmal aluI ban I 622 TTC AGC TAC TCT GGT GCC TTC AAG TGT CTG AGA GAC GGG GCT GGA GAC 190 Phe Ser Tyr Ser Gly Ala Phe Lys Cys Leu Arg Asp Gly Ala Gly Asp sau96I avall asul ppuMI hgiAI eco0109I bsp1286 mn l I mn l I 670 GTG GCT TTT ATC AGA GAG AGC ACA GTG TTT GAG GAC CTG TCA GAC GAG 206 Val Ala Phe Ile Arg Glu Ser Thr Val Phe Glu Asp Leu Ser Asp Glu 718 GCT GAA AGG GAC GAG TAT GAG TTA CTC TGC CCA GAC AAC ACT CGG AAG 222 Ala Glu Arg Asp Glu Tyr Glu Leu Leu Cys Pro Asp Asn Thr Arg Lys scrFI ncil mspI hpall caull xmal sau961 smal nlalv scrFI ncil avall caull aval asul sau96I ppuMI haellI nlaIV asul eco[]109I 766 CCA GTG GAC AAG TTC AAA GAC TGC CAT CTG GCC CGG GTC CCT TCT CAT 238 Pro Val Asp Lys Phe Lys Asp Cys His Leu Ala Arg Val Pro Ser His sfaNI fokl mboll mn l I dralll boll 814 GCC GTT GTG GCA CGA AGT GTG AAT GGC AAG GAG GAT GCC ATC TGG AAT 254 Ala Val Val Ala Arg Ser Val Asn Gly Lys Glu Asp Ala Ile Trp Asn scrFI ecoRII hph I bstNI 862 CTT CTC CGC CAG GCA CAG GAA AAG TTT GGA AAG GAC AAG TCA CCG AAA 270 Leu Leu Arg Gln Ala Gln Glu Lys Phe Gly Lys Asp Lys Ser Pro Lys

sau3AI mbol dpn I xholl aluI bstYI bstXI nlaIV bgllI 910 TTC CAG CTC TTT GGC TCC CCT AGT GGG CAG AAA GAT CTG CTG TTC AAG 286 Phe Gln Leu Phe Gly Ser Pro Ser Gly Gln Lys Asp Leu Leu Phe Lys nlaIV hgiCI pleI mn ll bsp1286 mn l I hinfl tagI ban I avaI 958 GAC TCT GCC ATT GGG TTT TCG AGG GTG CCC CCG AGG ATA GAT TCT GGG 302 Asp Ser Ala Ile Gly Phe Ser Arg Val Pro Pro Arg Ile Asp Ser Gly mspI styI hpaII nlaIV fokI rsal mn l I 1006 CTG TAC CTT GGC TCC GGC TAC TTC ACT GCC ATC CAG AAC TTG AGG AAA 318 Leu Tyr Leu Gly Ser Gly Tyr Phe Thr Ala Ile Gln Asn Leu Arg Lys msp I hpaII thal scrFI fnuDlI ncil bstUI mnlI fnu4HI hinPI bbvI cauII mniI hhaI 344 Ser Glu Glu Glu Val Ala Ala Arg Arg Ala Arg Val Val Trp Cys Ala hinPI mstI fspI fnu4HI aluI hhaI bstX1 alwNI bbvI bsrI 1102 GTG GGC GAG CAG GAG CTG CGC AAG TGT AAC CAG TGG AGT GGC TTG AGC 350 Val Gly Glu Gln Glu Leu Arg Lys Cys Asn Gln Trp Ser Gly Leu Ser fnu4H1 mn l I bbv I mn lI haeIII IMasd mn l I 1150 GAA GGC AGC GTG ACC TGC TCC TCG GCC TCC ACC ACA GAG GAC TGC ATC 366 Glu Gly Ser Val Thr Cys Ser Ser Ala Ser Thr Thr Glu Asp Cys Ile scrFI ecoRII bstXI bstNI aluI sfaNI nlaIII fokI 1198 GCC CTG GTG CTG AAA GGA GAA GCT GAT GCC ATG AGT TTG GAT GGA GGA 382 Ala Leu Val Leu Lys Gly Glu Ala Asp Ala Met Ser Leu Asp Gly Gly

FIG.-1D

nlaIII nlaIV scrF1 sphI hgiCI ecoRII rsal nspClx ban I bstNI 1246 TAT GTG TAC ACT GCA TGC AAA TGT GGT TTG GTG CCT GTC CTG GCA GAG 398 Tyr Val Tyr Thr Ala Cys Lys Cys Gly Leu Val Pro Val Leu Ala Glu sau3AI **MboI** don I alwI 1294 AAC TAC AAA TCC CAA CAA AGC AGT GAC CCT GAT CCT AAC TGT GTG GAT 414 Asn Tyr Lys Ser Gln Gln Ser Ser Asp Pro Asp Pro Asn Cys Val Asp sau3AI mbol ecoRV ecoNI don I 1342 AGA CCT GTG GAA GGA TAT CTT GCT GTG GCG GTG GTT AGG AGA TCA GAC 430 Arg Pro Val Glu Gly Tyr Leu Ala Val Ala Val Val Arg Arg Ser Asp scrfI ecoRII bstNI 1390 ACT AGC CTT ACC TGG AAC TCT GTG AAA GGC AAG AAG TCC TGC CAC ACC 446 Thr Ser Leu Thr Trp Asn Ser Val Lys Gly Lys Lys Ser Cys His Thr haeIII nlalll styl sau96I mboll ostI ncol asul 1438 GCC GTG GAC AGG ACT GCA GGC TGG AAT ATC CCC ATG GGC CTG CTC TTC 462 Ala Val Asp Arg Thr Ala Gly TrP Asn Ile Pro Met Gly Leu Leu Phe nlaIV hgiJII bsp1286 banll aluI bsp1286 sspl 1486 AAC CAG ACG GGC TCC TGC AAA TTT GAT GAA TAT TTC AGT CAA AGC TGT 478 Asn Gln Thr Gly Ser Cys Lys Phe Asp Glu Tyr Phe Ser Gln Ser Cys sau3AI Mbol Dpn I xhoII scrFl ecoRII bstYI hgiAI bstNI bglII bsp1286 avaI 1534 GCC CCT GGG TCT GAC CCG AGA TCT AAT CTC TGT GCT CTG TGT ATT GGC

484 Ala Pro Gly Ser Asp Pro Arg Ser Asn Leu Cys Ala Leu Cys Ile Gly

bsp1286 hphI 1582 GAC GAG CAG GGT GAG AAT AAG TGC GTG CCC AAC AGC AAT GAG AGA TAC 510 Asp Glu Gln Gly Glu Asn Lys Cys Val Pro Asn Ser Asn Glu Arg Tyr nlaIV hgICI ban1 scrF1 ecoRII mspI hpa II bstNI ddel bsml bsr I 1630 TAC GGC TAC ACT GGG GCT TTC CGG TGC CTG GCT GAG AAT GCT GGA GAC 526 Tyr Gly Tyr Thr Gly Ala Phe Arg Cys Leu Ala Glu Asn Ala Gly Asp 1678 GTT GCA TTT GTG AAA GAT GTC ACT GTC TTG CAG AAC ACT GAT GGA AAT 542 Val Ala Phe Val Lys Asp Val Thr Val Leu Gln Asn Thr Asp Gly Asn fnu4HI bbv I hinPI mnil nialli ddel aluI 1726 AAC AAT GAG GCA TGG GCT AAG GAT TTG AAG CTG GCA GAC TTT GCG CTG 558 Asn Asn Glu Ala Trp Ala Lys Asp Leu Lys Leu Ala Asp Phe Ala Leu fnu4HI taql mn l I bbv I mn l I dde I bglI 1774 CTG TGC CTC GAT GGC AAA CGG AAG CCT GTG ACT GAG GCT AGA AGC TGC 574 Leu Cys Leu Asp Gly Lys Arg Lys Pro Val Thr Glu Ala Arg Ser Cys sau96I nlaIV nialli styl haelll bsma I fokI hinfl nlaIII ncol asul 1822 CAT CTT GCC ATG GCC CCG AAT CAT GCC GTG GTG TCT CGG ATG GAT AAG 590 His Leu Ala Met Ala Pro Asn His Ala Val Val Ser Arg Met Asp Lys fnu4HI alwNI bbv1 ecoNI 187.0 GTG GAA CGC CTG AAA CAG GTG CTG CTC CAC CAA CAG GCT AAA TTT GGG 606 Val Glu Arg Leu Lys Gln Val Leu Leu His Gln Gln Ala Lys Phe Gly sau3AI mbol mspI hpall dpn I scrFl xholl ncil bstYI bsrl caull alwI 1919 AGA AAT GGA TCT GAC TGC CCG GAC AAG TTT TGC TTA TTC CAG TCT GAA 622 Arg Asn Gly Ser Asp Cys Pro Asp Lys Phe Cys Leu Phe Gln Ser Glu

FIG.-1F

```
haelll
                                                          hael
                                                          eael
                                                                       styl
                                                          cfrI
                                                                  pleI ncol
                                             ddeI
                                          dralll
                                                          ball
                                                                  hinfl
1966 ACC AAA AAC CTT CTG TTC AAT GAC AAC ACT GAG TGT CTG GCC AGA CTC
  638 Thr Lys Asn Leu Leu Phe Asn Asp Asn Thr Glu Cys Leu Ala Arg Leu
                                                sau96I
                                                avaII
                                                asuI
                                              nlaIV
                        ndel
                                     sspI
       nlaIII
 2014 CAT GGC AAA ACA ACA TAT GAA AAA TAT TTG GGA CCA CAG TAT GTC GCA
  654 His Gly Lys Thr Thr Tyr Glu Lys Tyr Leu Gly Pro Gln Tyr Val Ala
                                                           scrFI
                                                           ecoRII
                                                           bstNI
                                  hgIAI
                                  bsp1286
                                             mn l I
 2062 GGC ATT ACT AAT CGT AAA AAG TGC TCA ACC TCC CCC CTC CTG GAA GCC
  670 Gly Ile Thr Asn Leu Lys Lys Cys Ser Thr Ser Pro Leu Leu Glu Ala
                   dde I
                 mstII
                                                    sau96I
                 mn l I
                                                   haelll
                                            mboll
                 eco81I
                                                    asuI aluI
                                         mboll
           ecoRI bsu36I
 2110 TGT GAA TTC CTC AGG AAG TAA AACCGAAGAA GATGGCCCAG CTCCCCAAGA
  685 Cys Glu Phe Leu Arg Lys CX
                                                                styI
                                                             haeIII
                                                            sau96I
                                     mboll
                                              scrFI
                                                            asul
                                                           nlaIV
                                              ecoRI1
                                    earl
           dde I
                                                           eco[109] nlalV
                                              bstNI
                                  aluI
          mn l I
  2161 AAGCCTCAGC CATTCACTGC CCCCAGCTCT TCTCCCCAGG TGTGTTGGGG CCTTGGCTCC
                                fokl
       ecoNI
  2221 CCTGCTGAAG GTGGGGATTG CCCATCCATC TGCTTACAAT TCCCTGCTGT CGTCTTAGCA
  2281 AGAAGTAAAA TGAGAAATTT TGTTGATATT CAAAAAAAA
```

FIG.-1G

>LENGTH: 2319

rsal

OK to the substitute

```
nlaIII nlaIV scrFI
sphI hg:CI ecoRII
nspClx banI bstNI
```

1246 TAT GTG TAC ACT GCA TGC AAA TGT GGT TTG GTG CCT GTC CTG GCA GAG 398 Tyr Val Tyr Thr Ala Cys Lys Cys Gly Leu Val Pro Val Leu Ala Glu

sau3AI mboI dpnI alwI

1294 AAC TAC AAA TCC CAA CAA AGC AGT GAC CCT GAT CCT AAC TGT GTG GAT 414 Asn Tyr Lys Ser Gln Gln Ser Ser Asp Pro Asp Pro Asn Cys Val Asp

sau3AI mboI

ecoNI ecoRV dpn1

1342 AGA CCT GTG GAA GGA TAT CTT GCT GTG GCG GTG GTT AGG AGA TCA GAC 430 Arg Pro Val Glu Gly Tyr Leu Ala Val Ala Val Arg Arg Ser Asp

scrFI ecoRII bstNI

1390 ACT AGC CTT ACC TGG AAC TCT GTG AAA GGC AAG AAG TCC TGC CAC ACC 446 Thr Ser Leu Thr Trp Asn Ser Val Lys Gly Lys Lys Ser Cys His Thr

haeIII

nlaIII

styl sau96I mboll

pstI ncol asul earl
1438 GCC GTG GAC AGG ACT GCA GGC TGG AAT ATC CCC ATG GGC CTG CTC TTC
462 Ala Val Asp Arg Thr Ala Gly TrP Asn Ile Pro Met Gly Leu Leu Phe

nlaIV hgiJII bsp1286

ban II ssp 1 a lu I bsp 1286 1486 AAC CAG ACG GGC TCC TGC AAA TTT GAT GAA TAT TTC AGT CAA AGC TGT 478 Asn Gln Thr Gly Ser Cys Lys Phe Asp Glu Tyr Phe Ser Gln Ser Cys

sau3AI

mbol Dpn I

scrFI xholl ecoRII bstYl

hg:AI bsp1286

bstNI aval bgllI bsp1286
1534 GCC CCT GGG TCT GAC CCG AGA TCT AAT CTC TGT GCT CTG TGT ATT GGC
484 Ala Pro Gly Ser Asp Pro Arg Ser Asn Leu Cys Ala Leu Cys Ile Gly